

# Strong Secrecy and Stealth for Broadcast Channels with Confidential Messages

Igor Bjelakovic, Jafar Mohammadi, and Sławomir Stańczak

Fachgebiet Informationstheorie und theoretische Informationstechnik

Technische Universität Berlin, Einsteinufer 27, 10587 Berlin, Germany

Email: i.bjelakovic@tu-berlin.de, {jafar.mohammadi, slawomir.stanczak}@hhi.fraunhofer.de

**Abstract**—This paper extends the weak secrecy results of Liu et al. for broadcast channels with two confidential messages to strong secrecy. Our results are based on an extension of the techniques developed by Hou and Kramer on bounding Kullback-Leibler divergence in context of *resolvability* and *effective secrecy*.

**Index Terms**—Information theoretic security, resolvability, strong secrecy, broadcast channel.

## I. INTRODUCTION

Based on the pioneering work of Shannon [1], Wyner [2] determined the secrecy capacity of a class of wiretap channels. Wyner's work has been generalized by Csiszár and Körner [3] to the non-degraded broadcast channel with a single confidential message for one user and a common message intended for both users. The confidential message has to be kept secret from the other user, while both of them decode the common message. A variant of Csiszár and Körner's model with two confidential messages and no common message was first studied by Liu et al. in [4] for a discrete memoryless channel and later for the Gaussian case in [5]. An inner and outer bound on the secrecy region can be found in [4].

The secrecy criterion used in [2–4], is the *normalized* mutual information between the message and the output distribution of the user regarded as the eavesdropper. This security criterion, called weak secrecy, delivers a very restricted security against eavesdropping attacks as was shown by Maurer in [6]. An *unnormalized* definition of secrecy, called strong secrecy, was proposed in [6] and large parts of previous work was extended from weak to strong secrecy (e.g. [6], [7], [8], and, etc.). In this paper we extend the inner bound of [4] from weak to strong secrecy. Along the proof we extend the method of Hou and Kramer [9] based on even stronger notion of *effective secrecy* to this scenario. More precisely we prove, in addition to strong secrecy, a stealthy communication (i.e. the presence of meaningful communication is hidden) is possible.

### A. Notation

We use capital letters for random variables (RV). If  $X$  is a RV then  $x$  is used to refer to an observation of  $X$ . Sets are denoted by  $\mathcal{X}$  and  $\mathcal{P}(\mathcal{X})$  stands for the set of probability distributions defined on the (finite set)  $\mathcal{X}$ . For

the RVs  $X_1, \dots, X_K$  with values in  $\mathcal{X}_1, \dots, \mathcal{X}_K$  we write  $X_1 - X_2 - \dots - X_K$  if they form a Markov chain. The symbol  $\mathbf{x}$  stands for  $x^n := (x_1, \dots, x_n) \in \mathcal{X}^n$ . The mutual information between the RV's  $A$  and  $B$  is denoted by  $I(A, B)$ , while  $H(A)$  and  $H(A|B)$  are entropy of  $A$  and conditional entropy of  $A$  given  $B$ , respectively. Probability mass function of  $X$  is  $P_X(x)$  or in short  $P(x)$  while probability of an event  $E$  is denoted by  $\Pr(E)$ . Kullback-Leibler divergence of two probability distributions  $P, Q$  defined on set  $\mathcal{A}$  is given by,

$$D(P\|Q) := \begin{cases} \sum_{a \in \mathcal{A}} P(a) \log \frac{P(a)}{Q(a)} & \text{if } P \ll Q \\ +\infty & \text{if } P \not\ll Q \end{cases} \quad (1)$$

where  $P \ll Q$  means that  $P(a) = 0$  whenever  $Q(a) = 0$   $a \in \mathcal{A}$ . The typical set of sequences  $x^n \in \mathcal{X}^n$  for RV  $X$  and  $\epsilon > 0$  is denoted by  $\mathcal{T}_\epsilon^n(P_X)$  as defined in [10]. We will freely use the properties of typical sets from [10].

## II. SYSTEM MODEL

We consider a broadcast scenario consisting of a sender (S) and two receivers. We assume that all channels are discrete memoryless with finite input alphabet  $\mathcal{X}$ , and finite output alphabets  $\mathcal{Y}_1$  and  $\mathcal{Y}_2$ . The conditional probability distribution governing the discrete memoryless broadcast channel (DM-BCC) is given by

$$P(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{x}) = \prod_{i=1}^n P(y_{1i}, y_{2i} | x_i), \quad (2)$$

where,  $\mathbf{x} = x^n \in \mathcal{X}^n$ ,  $\mathbf{y}_t = y_t^n = (y_{t1}, \dots, y_{tn}) \in \mathcal{Y}_t^n$ , and  $t \in \{1, 2\}$ .

The stochastic encoder at the sender is defined to be,

$$f : \mathcal{M}_1 \times \mathcal{M}_2 \rightarrow \mathcal{P}(\mathcal{X}^n) \quad (3)$$

with,

$$\sum_{\mathbf{x} \in \mathcal{X}^n} f(\mathbf{x} | m_1, m_2) = 1 \quad \forall m_1 \in \mathcal{M}_1, m_2 \in \mathcal{M}_2 \quad (4)$$

where  $\mathcal{M}_1 := \{1, \dots, M_1\}$  and  $\mathcal{M}_2 := \{1, \dots, M_2\}$  are the message sets for receiver 1 and 2, respectively. The decoder at the  $t$ th node,  $t \in \{1, 2\}$ , is defined as

$$g_t : \mathcal{Y}_t^n \rightarrow \mathcal{M}_t. \quad (5)$$

*Definition 1 (Strong Secrecy [6]):* For every  $\epsilon > 0$  there is a non-negative integer  $N(\epsilon)$  such that for all  $n \geq N(\epsilon)$ ,

$$I(W_1; Y_2^n | W_2) \leq \epsilon \quad (6)$$

$$I(W_2; Y_1^n | W_1) \leq \epsilon \quad (7)$$

where the RVs  $W_1$  and  $W_2$  are distributed uniformly over  $\mathcal{M}_1$  and  $\mathcal{M}_2$ , respectively and the mutual information values are computed with respect to the distribution

$$P_{W_1 W_2 Y_1^n Y_2^n}(m_1, m_2, \mathbf{y}_1, \mathbf{y}_2) = \frac{1}{M_1} \frac{1}{M_2} \sum_{\mathbf{x} \in \mathcal{X}^n} f(x^n | m_1, m_2) P(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{x}),$$

with the stochastic encoder given in (3).

The probability of error at each node  $t \in \{1, 2\}$  is

$$P_{e,t}^n = \frac{1}{M_1 M_2} \sum_{(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2} P[g_t(y_t^n) \neq m_t | (m_1, m_2) \text{ is sent}],$$

*Definition 2:* A rate pair  $(R_1, R_2)$  is said to be achievable, if for every  $\epsilon' > 0$ ,  $\epsilon > 0$ ,  $\delta > 0$ , there exists a code  $(M_1, M_2, n, P_{e,1}^n, P_{e,2}^n)$  for sufficiently large  $n$ , in addition to (6) and (7), we have

$$P_{e,t}^n \leq \epsilon' \quad \text{and} \quad \frac{1}{n} \log M_t \geq R_t - \delta, \quad t \in \{1, 2\}.$$

*Remark 1:* The above definition provides a *strong* condition on security. More details are given in section III-C where we show that even stronger notion of secrecy based on stealth can be achieved.

### III. STRONG SECRECY FOR BROADCAST CHANNELS

In this section we present an achievable secure rate region with strong secrecy criterion for a DMBC. The following theorem summarizes our results:

*Theorem 1:* The rate pair  $(R_1, R_2)$  is achievable, in the sense of Definition 2, for DMBC with confidential messages and strong secrecy criterion, if

$$0 \leq R_1, 0 \leq R_2$$

$$R_1 < I(V_1; Y_1 | U) - I(V_1; Y_2 | V_2, U) - I(V_1; V_2 | U)$$

$$R_2 < I(V_2; Y_2 | U) - I(V_2; Y_1 | V_1, U) - I(V_1; V_2 | U)$$

where the information quantities are computed with respect to some probability distributions  $P_{U, V_1, V_2, X, Y_1, Y_2}$  such that Markov chain condition  $U - (V_1, V_2) - X - (Y_1, Y_2)$  holds and  $X$  and  $(Y_1, Y_2)$  are connected via the given broadcast channel. The auxiliary RVs  $U, V_1$  and  $V_2$  take values in finite sets  $\mathcal{U}, \mathcal{V}_1$  and  $\mathcal{V}_2$ .<sup>1</sup>

*Proof:* The proof, which consists of two parts, achievability and secrecy, unfolds in the following subsections. The achievability proof is based on techniques developed in [11, 12] and extends those devoted to secrecy developed in [9] for the wiretap channel to the present setting.

<sup>1</sup>The cardinalities of  $\mathcal{U}, \mathcal{V}_1, \mathcal{V}_2$  can be bounded by the Ahlswede-Kötner technique. We shall provide this in the journal version of this manuscript

#### A. Coding scheme

*Remark 2:* In order to simplify our notation we will drop the auxiliary random variable  $U$  in the following proof. It can be included into the proof via standard arguments [13, 14].

For each  $m_t \in \{1, \dots, 2^{nR_t}\}$ ,  $s_t \in \{1, \dots, 2^{nR'_t}\}$ , and  $k_t \in \{1, \dots, 2^{nR_{co}}\}$ ,  $t \in \{1, 2\}$ , we draw independently sequences  $\mathbf{v}_t(m_t, s_t, k_t)$  according to  $P^n(\mathbf{v}_t) = \prod_{i=1}^n P(v_{t,i})$ .

Let  $0 < \epsilon < \epsilon' < \epsilon''$ . For  $(m_1, m_2)$  and  $(s_1, s_2)$  find a pair  $(k_1, k_2)$  such that

$$(\mathbf{v}_1(m_1, s_1, k_1), \mathbf{v}_2(m_2, s_2, k_2)) \in \mathcal{T}_\epsilon^n(P_{V_1 V_2}).$$

If there is no such a pair choose  $(k_1, k_2) = (1, 1)$ . Then select a sequence  $\mathbf{x}(m_1, m_2, s_1, s_2) \in \mathcal{X}^n$  with

$$(\mathbf{v}_1(m_1, s_1, k_1), \mathbf{v}_2(m_2, s_2, k_2), \mathbf{x}(m_1, m_2, s_1, s_2)) \in \mathcal{T}_{\epsilon'}^n(P_{V_1 V_2 X}).$$

*Encoding:* To transmit  $(m_1, m_2)$  select uniformly at random a pair  $(s_1, s_2)$  and send  $\mathbf{x}(m_1, m_2, s_1, s_2)$ .

*Decoding:* Upon receiving  $\mathbf{y}_t$  decoder,  $t \in \{1, 2\}$ , declares  $(m_t, s_t)$  is sent if it is unique pair such that

$$(\mathbf{v}_t(m_t, s_t, k_t), \mathbf{y}_t) \in \mathcal{T}_{\epsilon''}^n(P_{V_t Y_t}) \quad (8)$$

#### B. Error Analysis

The error analysis is carried out using standard arguments. Using the mutual covering lemma, packing lemma, and the properties of the typical sequences (cf. for example [10]) we obtain

$$R_{co} > I(V_1; V_2) \quad \text{and} \quad R_1 + R'_1 + R_{co} < I(V_1; Y_1). \quad (9)$$

and

$$\lim_{n \rightarrow \infty} P_{e,t}^n = 0 \quad t \in \{1, 2\}$$

exponentially fast.

From (9) we have

$$R_1 + R'_1 < I(V_1; Y_1) - I(V_1; V_2) \quad (10)$$

and by symmetry,

$$R_2 + R'_2 < I(V_2; Y_2) - I(V_1; V_2). \quad (11)$$

The bounds on  $R'_1$  and  $R'_2$  are derived in the following subsection.

#### C. Strong Secrecy Criterion

*Remark 3:* For the secrecy analysis we drop the random variable  $U$  for the sake of notational simplicity. It can be introduced, if desired via standard arguments at the end of the proof [3, 9].

Before we further continue, we define a deterministic function  $\phi_\tau$  on pairs  $\{\mathbf{v}_1(m_1, s_1, k_1)\}_{k_1=1}^{2^{nR_{co}}} \times \{\mathbf{v}_2(m_2, s_2, k_2)\}_{k_2=1}^{2^{nR_{co}}}$  that returns a pair  $(k_1, k_2)$  such that

$$(\mathbf{v}_1(m_1, s_1, k_1), \mathbf{v}_2(m_2, s_2, k_2)) \in \mathcal{T}_\epsilon^n(P_{V_1, V_2}). \quad (12)$$

In the case that there are many such pairs, we choose one arbitrarily. However, if there are no such pairs, the function  $\phi_\tau$  returns (1,1). From now on, we denote the codeword pairs simply  $(\mathbf{v}_1(m_1, s_1), \mathbf{v}_2(m_2, s_2))$  based on selection function  $\phi_\tau$ .

We extend the framework in [9, 15] to find a condition that bounds

$$D(P_{W_1, Y_2|V_2} \| P_{W_1} Q_{Y_2|V_2}^n) \leq \xi \quad (13)$$

where the left hand side is known as the *effective secrecy*,  $\xi > 0$  is arbitrarily small and

$$P_{W_1, Y_2|V_2}(m_1, \mathbf{y}_2 | \mathbf{v}_2) := \frac{1}{M_1} \sum_{s_1=1}^{J_1} \frac{1}{J_1} P^n(\mathbf{y}_2 | \mathbf{v}_1(m_1, s_1), \mathbf{v}_2(m_2, s_2)),$$

and further  $Q_{Y_2|V_2}^n$  is distribution of  $Y_2^n$  given  $\mathbf{v}_2$  while no meaningful message is transmitted for  $W_1$ ,

$$Q_{Y_2|V_2}^n(\mathbf{y}_2 | \mathbf{v}_2) = \sum_{\mathbf{v}_1} P_{V_1}^n(\mathbf{v}_1) P_{Y_2|V_1, V_2}^n(\mathbf{y}_2 | \mathbf{v}_1, \mathbf{v}_2) \quad (14)$$

and finally,  $P_{W_1}$  is distribution of  $W_1$ . We can further expand (13) as

$$\begin{aligned} D(P_{W_1, Y_2|V_2} \| P_{W_1} Q_{Y_2|V_2}^n) &= \sum_{m_1, \mathbf{y}_2} P_{W_1, Y_2|V_2}(m_1, \mathbf{y}_2 | \mathbf{v}_2) \\ &\quad \log \left( \frac{P_{W_1, Y_2|V_2}(m_1, \mathbf{y}_2 | \mathbf{v}_2)}{Q_{Y_2|V_2}^n(\mathbf{y}_2 | \mathbf{v}_2) P_{W_1}(m_1)} \right) \\ &= \sum_{m_1, \mathbf{y}_2} P_{W_1, Y_2|V_2}(m_1, \mathbf{y}_2 | \mathbf{v}_2) \\ &\quad \log \left( \frac{P_{W_1, Y_2|V_2}(m_1, \mathbf{y}_2 | \mathbf{v}_2)}{Q_{Y_2|V_2}^n(\mathbf{y}_2 | \mathbf{v}_2) P_{W_1}(m_1)} \cdot \frac{P_{Y_2|V_2}(\mathbf{y}_2 | \mathbf{v}_2)}{P_{Y_2|V_2}(\mathbf{y}_2 | \mathbf{v}_2)} \right) \\ &= \sum_{m_1, \mathbf{y}_2} P_{W_1, Y_2|V_2}(m_1, \mathbf{y}_2 | \mathbf{v}_2) \\ &\quad \log \left( \frac{P_{W_1, Y_2|V_2}(m_1, \mathbf{y}_2 | \mathbf{v}_2)}{P_{W_1}(m_1) P_{Y_2|V_2}(\mathbf{y}_2 | \mathbf{v}_2)} \cdot \frac{P_{Y_2|V_2}(\mathbf{y}_2 | \mathbf{v}_2)}{Q_{Y_2|V_2}^n(\mathbf{y}_2 | \mathbf{v}_2)} \right) \\ &\stackrel{(\alpha)}{=} \sum_{m_1, \mathbf{y}_2} P_{W_1, Y_2|V_2}(m_1, \mathbf{y}_2 | \mathbf{v}_2) \log \left( \frac{P_{W_1, Y_2|V_2}(m_1, \mathbf{y}_2 | \mathbf{v}_2)}{P_{W_1}(m_1) P_{Y_2|V_2}(\mathbf{y}_2 | \mathbf{v}_2)} \right) \\ &\quad + \sum_{\mathbf{y}_2} P_{Y_2|V_2}(\mathbf{y}_2 | \mathbf{v}_2) \log \left( \frac{P_{Y_2|V_2}(\mathbf{y}_2 | \mathbf{v}_2)}{Q_{Y_2|V_2}^n(\mathbf{y}_2 | \mathbf{v}_2)} \right) \\ &= I(W_1; Y_2^n | \mathbf{v}_2) + D(P_{Y_2|V_2} \| Q_{Y_2|V_2}^n), \end{aligned} \quad (15)$$

in (α) we used that  $P_{Y_2|V_2}(\mathbf{y}_2 | \mathbf{v}_2) = \sum_{m_1} P(m_1, \mathbf{y}_2 | \mathbf{v}_2)$ .

*Remark 4:* The expansion in (15), reveals that (13) addresses not only the strong secrecy notation by bounding  $I(W_1; Y_2^n | \mathbf{v}_2)$ , but also bounds  $D(P_{Y_2|V_2} \| Q_{Y_2|V_2}^n)$  which corresponds to *stealth* of the communication.

Based on chain rule for informational divergence we have,

$$\begin{aligned} D(P_{W_1, Y_2|V_2} \| P_{W_1} Q_{Y_2|V_2}^n) &= \\ D(P_{W_1} \| P_{W_1}) + D(P_{Y_2|V_2, W_1} \| Q_{Y_2|V_2}^n | P_{W_1}), \end{aligned} \quad (16)$$

where, according to the definition of  $D(\cdot \| \cdot)$  in (1), the first term above equals to zero, thus we proceed with bounding  $D(P_{Y_2|V_2, W_1} \| Q_{Y_2|V_2}^n | P_{W_1})$ . The following lemma provides an upperbound on  $D(P \| Q)$ , which is useful for the rest of the proof.

*Lemma 1:* For probability distributions  $P$  and  $Q$ , defined on a finite set  $A$ , with  $P \ll Q$ , we have,

$$D(P \| Q) \leq \log \frac{1}{\pi_Q} \quad (17)$$

where,  $\pi_Q = \min\{Q(a) : Q(a) > 0\}$ .

Taking expectation with respect to  $(V_1^n, V_2^n)$  of  $\mathbb{1}(\phi_\tau = (1, 1)) D(P_{Y_2|V_2, W_1} \| Q_{Y_2|V_2}^n | P_{W_1})$  yields,

$$\begin{aligned} E[\mathbb{1}(\phi_\tau = (1, 1)) D(P_{Y_2|V_2, W_1} \| Q_{Y_2|V_2}^n | P_{W_1})] &= \Pr(\phi_\tau = (1, 1)) \log \frac{1}{(\pi_{Q_{Y_2|V_2}})^n} \\ &= \Pr(\phi_\tau = (1, 1)) (-n) \log(\pi_{Q_{Y_2|V_2}}) \\ &\leq 2^{-n\alpha} \end{aligned} \quad (18)$$

for all sufficiently large  $n$ , where  $\mathbb{1}(\cdot)$  is an indicator function and we apply lemma 1 in (18). The last inequality comes from the mutual covering lemma (cf. [10] for example), where  $\alpha > 0$  is a constant independent of  $n$ . With (16) and (19), therefore, we have for all sufficiently large  $n$

$$\begin{aligned} E[D(P_{Y_2|W_1, V_2} \| Q_{Y_2|V_2} | P_{W_1})] &\leq \\ E[\mathbb{1}(\phi_\tau \neq (1, 1)) D(P_{Y_2|V_2, W_1} \| Q_{Y_2|V_2} | P_{W_1}) + 2^{-n\alpha}] & \end{aligned} \quad (20)$$

Hence, in the following we focus on bounding the first term in the right hand side of (20).

We take the expectation over the  $Y_2^n$ ,  $V_1$ ,  $W_1$  and  $S_1$ ; we obtain

$$\begin{aligned} E[D(P_{Y_2|W_1, V_2} \| Q_{Y_2|V_2}^n | P_{W_1})] &= E \left[ \log \frac{\sum_{s_1=1}^{J_1} P(Y_2^n | V_1^n(W_1, s_1), \mathbf{v}_2)}{J_1 Q_{Y_2|V_2}^n(Y_2^n | \mathbf{v}_2)} \right] \\ &= \sum_{m_1, s_1} \frac{1}{J_1 M_1} E \left[ \log \frac{\sum_{i=1}^{J_1} P(Y_2^n | V_1^n(m_1, i), \mathbf{v}_2)}{J_1 Q_{Y_2|V_2}^n(Y_2^n | \mathbf{v}_2)} \right] \\ &\quad \left[ W_1 = m_1, S_1 = s_1 \right] \\ &\leq \sum_{m_1, s_1} \frac{1}{J_1 M_1} E \left[ \log \frac{P(Y_2^n | V_1^n(m_1, s_1), \mathbf{v}_2)}{J_1 Q_{Y_2|V_2}^n(Y_2^n | \mathbf{v}_2)} \right. \\ &\quad \left. + \frac{J_1 - 1}{J_1} \right] \left[ W_1 = m_1, S_1 = s_1 \right] \end{aligned} \quad (21)$$

$$\begin{aligned} &\leq \sum_{m_1, s_1} \frac{1}{J_1 M_1} E \left[ \log \frac{P(Y_2^n | V_1^n(m_1, s_1), \mathbf{v}_2)}{J_1 Q_{Y_2|V_2}^n(Y_2^n | \mathbf{v}_2)} + 1 \right] \\ &\quad \left[ W_1 = m_1, S_1 = s_1 \right] \\ &= E \left[ \log \left( \frac{P(Y_2^n | V_1^n, \mathbf{v}_2)}{J_1 Q_{Y_2|V_2}^n(Y_2^n | \mathbf{v}_2)} + 1 \right) \right], \end{aligned} \quad (22)$$

where, in (21) we used Jensen's inequality applied to the part of the expectation over  $V_1^n(m_1, i)$  for  $i \neq s_1$ .<sup>2</sup>

Before proceeding with the bounds on the informational divergence, we introduce the following lemma.

**Lemma 2:** Let  $P_{V_1, V_2, Y_2}(y_2, v_2, v_1) = P_{V_1, V_2}(v_1, v_2)P_{Y_2|V_1, V_2}(y_2|v_1, v_2)$  be a probability distribution on  $\mathcal{Y}_2 \times \mathcal{V}_1 \times \mathcal{V}_2$ . For  $\epsilon \in (0, \frac{1}{2})$  and distribution

$$Q(y_2|v_2) := \sum_{v_1} P_{V_1}(v_1)P_{Y_2|V_1, V_2}(y_2|v_1, v_2) \quad (23)$$

it holds for  $(\mathbf{v}_2, \mathbf{y}_2) = \mathcal{T}_\epsilon^n(P_{V_2, Y_2})$  that

$$Q^n(\mathbf{y}_2|\mathbf{v}_2) \geq 2^{-n(H(Y_2|V_2) + I(V_1; V_2) + \delta(\epsilon))} \quad (24)$$

with  $\delta(\epsilon) > 0$  and  $\lim_{\epsilon \rightarrow 0} \delta(\epsilon) = 0$ .

*Proof:* According to Lemma 2.6 in [13] we have for all  $(Y_2, V_2)$ ,

$$Q^n(y_2|v_2) = 2^{-n(H(P_{V_2, Y_2}^e) - H(P_{V_2}^e) + D(P_{V_2, Y_2}^e \| P_{V_2}^e Q))} \quad (25)$$

where,

- 1)  $P_{V_2, Y_2}^e$  denotes the empirical distribution or type of the sequence pair  $(\mathbf{v}_2, \mathbf{y}_2)$
- 2)  $P_{V_2}^e$  is the empirical distribution generated by the sequence  $\mathbf{v}_2$
- 3) the distribution  $P_{V_2}^e Q$  is the joint distribution computed with respect to  $P_{V_2}^e$  and  $Q$

Uniform continuity of the entropy (Lemma 2.7 in [13]),  $(\mathbf{v}_2, \mathbf{y}_2) = \mathcal{T}_\epsilon^n(P_{V_2, Y_2})$ , and the fact that  $\epsilon \in (0, \frac{1}{2})$ , yield

$$|H(P_{V_2, Y_2}^e) - H(P_{V_2}^e) - H(Y_2|V_2)| \leq \delta_1(\epsilon), \quad (26)$$

where,  $\delta_1(\epsilon) > 0$  with  $\lim_{\epsilon \rightarrow 0} \delta_1(\epsilon) = 0$ . Moreover,

$$|D(P_{V_2, Y_2}^e \| P_{V_2}^e Q) - D(P_{V_2, Y_2} \| P_{V_2} Q)| \leq \delta_2(\epsilon), \quad (27)$$

with  $\delta_2(\epsilon) > 0$  and  $\lim_{\epsilon \rightarrow 0} \delta_2(\epsilon) = 0$ .

Inequality (27) can be seen as follows: In a first step we show that  $P_{V_2, Y_2}^e \ll P_{V_2}^e Q$ . To this end we assume

$$P_{V_2}^e Q(v_2, y_2) = P_{V_2}^e(v_2)Q(y_2|v_2) = 0.$$

Then either  $P_{V_2}^e(v_2) = 0$  and then  $P_{V_2, Y_2}^e(v_2, y_2) = 0$  automatically. Or  $Q(y_2|v_2) = 0$  and then

$$\begin{aligned} 0 &= Q(y_2|v_2) \\ &= \sum_{v_1} P_{V_1}(v_1)P_{Y_2|V_1, V_2}(y_2|v_1, v_2) \\ &\geq \sum_{v_1} P_{V_1, V_2}(v_1, v_2)P_{Y_2|V_1, V_2}(y_2|v_1, v_2) \\ &= \sum_{v_1} P_{V_1, V_2, Y_2}(v_1, v_2, y_2) \\ &= P_{V_2, Y_2}(v_2, y_2), \end{aligned} \quad (28)$$

which, again, implies  $P_{V_2, Y_2}^e(v_2, y_2) = 0$  by definition of typical sequences, and thus  $P_{V_2, Y_2}^e \ll P_{V_2}^e Q$ . Since  $P_{V_2, Y_2}^e \ll$

$P_{V_2}^e Q$  we have

$$\begin{aligned} D(P_{V_2, Y_2}^e \| P_{V_2}^e Q) &= \sum_{v_2, y_2} P_{V_2, Y_2}^e(v_2, y_2) \log P_{V_2, Y_2}^e(v_2, y_2) \\ &\quad - \sum_{v_2, y_2} P_{V_2, Y_2}^e(v_2, y_2) \log P_{V_2}^e(v_2) \\ &\quad - \sum_{v_2, y_2} P_{V_2, Y_2}^e(v_2, y_2) \log Q(y_2|v_2) \end{aligned} \quad (29)$$

and applying again the uniform continuity of entropy along with the properties of typical sequences we obtain (27).

To proceed with the proof we need to show,

$$D(P_{V_2, Y_2} \| P_{V_2} Q) \leq I(V_1, V_2). \quad (30)$$

This inequality can be derived as follows,

$$\begin{aligned} D(P_{V_2, Y_2} \| P_{V_2} Q) &= \sum_{a, b} P_{V_2, Y_2}(a, b) \log \frac{P_{V_2, Y_2}(a, b)}{P_{V_2}(a) \sum_c P_{V_1}(c) P_{Y_2|V_1, V_2}(b|c, a)} \\ &= \sum_{c, a, b} P_{V_1, V_2, Y_2}(c, a, b) \log \frac{\sum_c P_{V_1, V_2, Y_2}(c, a, b)}{\sum_c P_{V_2}(a) P_{V_1}(c) P_{Y_2|V_1, V_2}(b|c, a)} \\ &\leq \sum_{c, a, b} P_{V_1, V_2, Y_2}(c, a, b) \log \frac{P_{V_1, V_2, Y_2}(c, a, b)}{P_{V_2}(a) P_{V_1}(c) P_{Y_2|V_1, V_2}(b|c, a)} \\ &= \sum_{c, a, b} P_{V_1, V_2, Y_2}(c, a, b) \log \frac{P_{V_1, V_2}(c, a) P_{Y_2|V_1, V_2}(b|c, a)}{P_{V_2}(a) P_{V_1}(c) P_{Y_2|V_1, V_2}(b|c, a)} \\ &= \sum_{c, a, b} P_{V_1, V_2, Y_2}(c, a, b) \log \frac{P_{V_1, V_2}(c, a)}{P_{V_2}(a) P_{V_1}(c)} \\ &= \sum_{c, a} P_{V_1, V_2}(c, a) \log \frac{P_{V_1, V_2}(c, a)}{P_{V_2}(a) P_{V_1}(c)} \\ &= I(V_1, V_2), \end{aligned} \quad (31)$$

where, (31) is by the Log-Sum-Inequality (Lemma 3.1 in [13]).

Combining (26), (27), and (32) leads to the claim of the lemma.  $\blacksquare$

The expression (22) can be split up as follows: (Recall that the expectation is taken over those codewords with  $\phi_\tau \neq (1, 1)$ )

$$\begin{aligned} &\sum_{(\mathbf{v}_1, \mathbf{v}_2) \in \mathcal{T}_\epsilon^n} P_{V_1}^n(\mathbf{v}_1) \sum_{\mathbf{y}_2} P_{Y_2|V_1, V_2}^n(\mathbf{y}_2|\mathbf{v}_1, \mathbf{v}_2) \\ &\log \left( \frac{P_{Y_2|\mathbf{v}_2, \mathbf{v}_1}^n(\mathbf{y}_2|\mathbf{v}_2, \mathbf{v}_1)}{J_1 Q_{Y_2|\mathbf{v}_2}^n(\mathbf{y}_2|\mathbf{v}_2)} + 1 \right) = e_1 + e_2 \end{aligned}$$

<sup>2</sup>A detailed proof of this derivation is due to journal version of this work

where,

$$e_1 := \sum_{\substack{\mathbf{v}_1 \\ (\mathbf{v}_1, \mathbf{v}_2) \in \mathcal{T}_\epsilon^n}} P_{V_1}^n(\mathbf{v}_1) \sum_{\substack{\mathbf{y}_2 \\ (\mathbf{v}_1, \mathbf{v}_2, \mathbf{y}_2) \notin \mathcal{T}_\epsilon^n}} P_{Y_2|V_1, V_2}^n(\mathbf{y}_2|\mathbf{v}_1, \mathbf{v}_2) \log \left( \frac{P^n(\mathbf{y}_2|\mathbf{v}_2, \mathbf{v}_1)}{J_1 Q^n(\mathbf{y}_2|\mathbf{v}_2)} + 1 \right) \quad (33)$$

$$e_2 := \sum_{\substack{\mathbf{v}_1 \\ (\mathbf{v}_1, \mathbf{v}_2) \in \mathcal{T}_\epsilon^n}} P_{V_1}^n(\mathbf{v}_1) \sum_{\substack{\mathbf{y}_2 \\ (\mathbf{v}_1, \mathbf{v}_2, \mathbf{y}_2) \in \mathcal{T}_\epsilon^n}} P_{Y_2|V_1, V_2}^n(\mathbf{y}_2|\mathbf{v}_1, \mathbf{v}_2) \log \left( \frac{P^n(\mathbf{y}_2|\mathbf{v}_2, \mathbf{v}_1)}{J_1 Q^n(\mathbf{y}_2|\mathbf{v}_2)} + 1 \right) \quad (34)$$

We upperbound  $e_1$  in (33) as

$$\begin{aligned} e_1 &\leq \sum_{\substack{\mathbf{v}_1 \\ (\mathbf{v}_1, \mathbf{v}_2) \in \mathcal{T}_\epsilon^n}} P_{V_1}^n(\mathbf{v}_1) \sum_{\substack{\mathbf{y}_2 \\ (\mathbf{v}_1, \mathbf{v}_2, \mathbf{y}_2) \notin \mathcal{T}_\epsilon^n}} P_{Y_2|V_1, V_2}^n(\mathbf{y}_2|\mathbf{v}_1, \mathbf{v}_2) \log \left( \left( \frac{1}{\pi_{Y_2|V_2}} \right)^n + 1 \right) \\ &\leq 2|\mathcal{Y}_2||\mathcal{V}_1||\mathcal{V}_2|2^{-2n\epsilon^2\pi_{Y_2|V_1, V_2}^2}(-n) \log(\pi_{Y_2|V_1}). \quad (35) \end{aligned}$$

where  $\pi_{Y_2|V_2}$  and  $\pi_{Y_2|V_1, V_2}$  are derived from (17). The (35) implies that as  $n \rightarrow \infty$ ,  $e_1 \rightarrow 0$ .

On the other hand, to upper bound  $e_2$  in (34) we need to use the Lemma 2, since,  $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{y}_2) \in \mathcal{T}_\epsilon^n(P_{V_1 V_2 Y_2})$  defined on  $P_{V_1, V_2, Y_2}$ . Therefore we have,

$$\begin{aligned} e_2 &\leq \sum_{\substack{\mathbf{v}_1 \\ (\mathbf{v}_1, \mathbf{v}_2) \in \mathcal{T}_\epsilon^n}} P_{V_1}^n(\mathbf{v}_1) \log \left( \frac{2^{-n(H(Y_2|V_1, V_2) - \delta_1(\epsilon))}}{J_1 2^{-n(H(Y_2|V_2) + I(V_1; V_2) + \delta_2(\epsilon))}} + 1 \right) \\ &\leq \sum_{\substack{\mathbf{v}_1 \\ (\mathbf{v}_1, \mathbf{v}_2) \in \mathcal{T}_\epsilon^n}} 2^{-n(H(V_1) + I(V_1; V_2) - \delta_3(\epsilon))} \log \left( \frac{2^{-n(H(Y_2|V_1, V_2) - \delta_1(\epsilon))}}{J_1 2^{-n(H(Y_2|V_2) + I(V_1; V_2) + \delta_2(\epsilon))}} + 1 \right) \\ &\leq 2^{n(H(V_1) + \delta_4(\epsilon))} 2^{-n(H(V_1) + I(V_1; V_2) - \delta_3(\epsilon))} \log \left( \frac{2^{-n(H(Y_2|V_1, V_2) - \delta_1(\epsilon))}}{J_1 2^{-n(H(Y_2|V_2) + I(V_1; V_2) + \delta_2(\epsilon))}} + 1 \right) \\ &\leq 2^{-n(I(V_1; V_2) - \delta_3(\epsilon) - \delta_4(\epsilon))} \log \left( \frac{2^{-n(H(Y_2|V_1, V_2) - \delta_1(\epsilon))}}{J_1 2^{-n(H(Y_2|V_2) + I(V_1; V_2) + \delta_2(\epsilon))}} + 1 \right) \\ &\leq 2^{-n(I(V_1; V_2) - \delta_3(\epsilon) - \delta_4(\epsilon))} \log \left( \frac{2^{-n(H(Y_2|V_1, V_2) - \delta_1(\epsilon))}}{J_1 2^{-n(H(Y_2|V_2) + I(V_1; V_2) + \delta_2(\epsilon))}} + 1 \right) \\ &\leq 2^{-n(I(V_1; V_2) - \delta_3(\epsilon) - \delta_4(\epsilon))} \log \left( 2^{-n(R'_1 + H(Y_2|V_1, V_2) - H(Y_2|V_2) - I(V_1; V_2) - \delta_1(\epsilon) - \delta_2(\epsilon))} + 1 \right) \\ &\leq \log(e) 2^{-n(R'_1 - I(Y_2; V_1|V_2) - \delta_5(\epsilon))}, \end{aligned}$$

where,  $\delta_5(\epsilon) := \delta_1(\epsilon) + \delta_2(\epsilon) + \delta_3(\epsilon) + \delta_4(\epsilon)$ . We have to impose  $R'_1 > I(Y_2; V_1|V_2) + \delta_5(\epsilon)$  in order to assure  $e_2 \rightarrow 0$  with  $n \rightarrow \infty$ . This condition, thus, evokes strong secrecy.

Combining  $R'_1 > I(Y_2; V_1|V_2)$ , and symmetrically  $R'_2 > I(Y_1; V_2|V_1)$ , (10) and (11), we obtain the rate region as in the theorem 1. ■

#### IV. ACKNOWLEDGEMENT

This work is supported by the German Research Foundation (DFG) under Grant STA 864/7-1.

#### REFERENCES

- [1] C. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.
- [2] A. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, pp. 1355–1387, Oct. 1975.
- [3] I. Csiszar and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.
- [4] R. Liu, I. Maric, P. Spasojevic, and R. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *Information Theory, IEEE Transactions on*, vol. 54, no. 6, pp. 2493–2507, June 2008.
- [5] R. Liu and H. Poor, "Secrecy capacity region of a multiple-antenna gaussian broadcast channel with confidential messages," *Information Theory, IEEE Transactions on*, vol. 55, no. 3, pp. 1235–1249, March 2009.
- [6] U. Maurer, "The strong secret key rate of discrete random triples," *Communications and Cryptography, -Two Sides of One Tapestry, Kluwer Academic Publishers*, vol. 276, pp. 271–285, 1994.
- [7] I. Csiszar, "Almost independence and secrecy capacity," *Probl. Peredachi Inf.*, vol. 32:1, pp. 48–57, 1996.
- [8] M. Bloch and J. Laneman, "Strong secrecy from channel resolvability," *Information Theory, IEEE Transactions on*, vol. 59, no. 12, pp. 8077–8098, Dec 2013.
- [9] J. Hou and G. Kramer, "Effective secrecy: Reliability, confusion and stealth," in *Information Theory (ISIT), 2014 IEEE International Symposium on*, June 2014, pp. 601–605.
- [10] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. New York, NY, USA: Cambridge University Press, 2012.
- [11] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.
- [12] S. I. Gel'fand and M. S. Pinsker, "Coding for Channel with Random Parameters," *Problems of Control Theory*, vol. 9, no. 1, pp. 19–31, 1980.
- [13] I. Csiszar and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Orlando, FL, USA: Academic Press, Inc., 1982.
- [14] A. El Gamal and E. van der Meulen, "A proof of marton's coding theorem for the discrete memoryless broadcast channel (corresp.)," *Information Theory, IEEE Transactions on*, vol. 27, no. 1, pp. 120–122, Jan 1981.
- [15] J. Hou and G. Kramer, "Informational divergence approximations to product distributions," in *Information Theory (CWIT), 2013 13th Canadian Workshop on*, June 2013, pp. 76–81.